高通量转录组测序技术及其在鳞翅目昆虫上的应用

杨 帆[1]，黄立华[2]，张爱兵[1,*]

（1. 首都师范大学生命科学学院，北京 100048；2. 华南师范大学生命科学学院，广州 510631）

摘要：转录组是指组织或者细胞在某一特定状态下转录出来的所有 RNA 的集合。高通量第 2 代测序技术使转录组学的研究模式发生了巨大的改变，所衍生出的转录组测序迅速成为研究非模式生物的先进技术。转录组测序能够在整体水平上探究细胞内基因表达的种类和数量，揭示在特定条件下机体生理生化发生过程以及其中的分子机理。本文简要阐述了转录组测序技术的基本概念、技术流程与原理，详细介绍了转录组测序在解决鳞翅目昆虫的分类、毒理、发育、与寄主互作以及非编码 RNA 调控等问题上做出的贡献，并对该技术现存的困难进行了系统的阐述并对其未来的发展趋势作出了简要的预测与剖析。

关键词：转录组；转录组测序；高通量测序；数据组装；测序流程；鳞翅目

中图分类号：Q966　　**文献标识码**：A　　**文章编号**：0454-6296(2014)08-0991-10

High-throughput transcriptome sequencing technology and its applications in Lepidoptera

YANG Fan[1], HUANG Li-Hua[2], ZHANG Ai-Bing[1,*] (1. College of Life Sciences, Capital Normal University, Beijing 100048, China; 2. School of Life Sciences, South China Normal University, Guangzhou 510631, China)

Abstract: Transcriptome is the complete set of RNA transcripts in a specific type of cell or tissue under a certain stage or physiological condition. Next generation high-throughput sequencing technology has changed the study patterns of transcriptomics tremendously, from which RNA-sequencing (RNA-Seq) is derived. Next generation sequencing is rapidly becoming an advanced technology to explore genetic outlines of non-model organisms. It helps to elucidate the types and number of expressed genes comprehensively, which reveals the molecular mechanisms of physiological and biochemical processes in organisms under certain conditions. This article introduced briefly the basic concepts of transcriptome sequencing technology and technical processes, elaborated the contribution of this technology to solve problems about taxonomy, toxicology, growth and development, insect-plant interactions, and even non-coding RNA in Lepidoptera, and elucidated the current technical difficulties and the future trends of this technology.

Key words: Transcriptome; transcriptome sequencing; high-throughput sequencing; data assembly; sequencing processes; Lepidoptera

转录组指在某一特定的生长阶段或生理条件下，细胞内所有转录产物的集合。它包括信使 RNA (mRNA)、核糖体 RNA (rRNA)、转运 RNA (tRNA) 及非编码 RNA。如果能够在最大程度上了解这些遗传物质的种类、数量、转录方式以及剪接模式等信息，并且定量地描述出在不同的特定条件下这些信息的区别，就可以直观地将某物种基因型和表型建立联系，从而极大地促进我们对细胞内基因组成与功能(David *et al.*, 2006; Zhang *et al.*, 2010a)、生长发育(Van Belleghem *et al.*, 2012)、毒理生化(Zhu *et al.*, 2011)以及系统发育(Wang *et al.*, 2010)等方面的研究。

1 转录组研究的主要技术手段

转录组的研究目前主要基于两大类技术手段：

基金项目：国家自然科学基金项目(31272340,30700641)；北京市人才强教项目(PHR201107120)；中华人民共和国科学技术部科技基础性工作专项重点项目(2012FY110800)

作者简介：杨帆，女，1987 年生，北京人，博士研究生，研究方向为昆虫遗传多样性及分子生物学，E-mail: yangfan3095861@sina.com

* 通讯作者 Corresponding author, E-mail: zhangab2008@cnu.edu.cn

收稿日期 Received: 2014-02-24；接受日期 Accepted: 2014-06-20

分子杂交与测序技术。分子杂交属于高通量研究中相对便宜的实验手段，主要利用 DNA 碱基互补的原理，将提前设计好的探针与具有荧光标记的 cDNA 进行结合，利用专门的仪器检测荧光亮度，并将其转换成数字加以比较（Bertone *et al.*，2004；Cheng *et al.*，2005；David *et al.*，2006）。例如 Huang 等（2009）就利用家蚕 *Bombyx mori* 的 cDNA 芯片与感染黑胸败血芽孢杆菌 *Bacillus bombysептieus* 的家蚕的转录组杂交，通过对杂交信号的分析探究家蚕的感染后转录组反应。然而分子杂交技术有其自身的缺陷，例如比较不同实验的杂交信号结果需要复杂的标准化过程去除背景“噪音”；而且探针与 cDNA 序列之间可能发生假阳性杂交，影响实验结果（Okoniewski and Miller，2006；Royce *et al.*，2007）。最重要的一点不足是，分子杂交技术非常依赖现有的基因组信息，因而将研究对象限制在了那些已经拥有参考序列的物种上面，而很多非模式物种并没有基因组全序列的信息，因而无法通过此方法对其转录组进行研究（Yang *et al.*，2013）。

不同于分子杂交技术对现有参考序列的依赖，测序技术的兴起可以让人们直接获悉 cDNA 序列的组成。早期所使用的 Sanger 测序（Boguski *et al.*，1994）又称为第 1 代测序，是一种非常实用的方法，但其通量低、价格高而且几乎无法定量等弊端使它无法满足更加精细的组学类研究的要求。基因表达系列分析（Velculescu *et al.*，1995；Harbers and Carninci，2005）可以通过限制性酶切基因组获得成千上万个表达序列标签，快速详细地分析这些短 cDNA（10～14 bp）可以探寻序列的表达丰富度，了解特定组织或细胞类型中基因诸多的表达状态，可以说是一种可量化的高通量研究手段，然而这些标签在进行 PCR 扩增和连接后仍然依赖于传统的一代测序，因此费用仍然是其限制因素之一。高通量测序技术（又称第 2 代测序技术）的出现和发展则突破了这些局限，为组学的研究打开了新局面。

2 转录组测序基本原理

第 2 代高通量 DNA 测序技术因为其单次运行（run）产出序列数据量大而得名（Ansorge，2009），其获得单个碱基的测序费用远远低于第 1 代测序。近年来它的高速发展使得在很短时间内可以测出众多生物物种的基因组序列，为整个生物学的研究带来了极大的突破（Yue *et al.*，2012）。第 2 代高通量测序在准确率、延长度和性价比三方面都明显优于一代测序，因此又被广泛应用于转录组的研究中，从而衍生出转录组测序（RNA sequencing，RNA-Seq）一词，即利用第 2 代高通量测序技术全面快速地获取某一物种特定器官或组织在某一特定状态下所有的转录本信息（主要包括 mRNA 和非编码 RNA）。所获得的海量数据经过专业生物信息软件的比对、组装等分析后，就可以还原出不同时空条件下不同组织或细胞中基因表达的各类特征（图 1）。

3 转录组测序主要的技术平台与基本测序流程

主流的转录组测序技术主要有 3 种，分别为 2005 年兴起的罗氏 454 焦磷酸测序技术（Emrich *et al.*，2007）、2006 年兴起的 Illumina/Solexa 聚合酶合成测序（Morin *et al.*，2008）以及 2007 年兴起的 ABI/SOLiD 连接酶测序技术（Cloonan *et al.*，2008）。罗氏 454 焦磷酸测序法获得的读长（reads）最长，约为 600～1 000 bp，但其通量较低，单次运行（run）产出数据为 0.5～1 Gb；Illumina/Solexa 利用单分子簇的边合成边测序技术，获得的读长通常为 100 bp 左右，但是其测序通量大，如 Hiseq2000 机型的单次产出数据量可达到 600 G；ABI/SOLiD 利用 DNA 磁珠获得的读长最短（50 bp），但是其双碱基编码和克隆链接校正系统可以有效地降低测序错误率，因此 SOLiD 特别适用于具有高质量参考基因组序列物种的重测序（Yue *et al.*，2012）。虽然 3 种第 2 代测序技术在转录组的测序原理、数据量产出以及单 Run 运行成本上均有不同之处（Glenn，2011），但是基本的测序流程比较相似：1）研究对象的总 RNA 提取；2）总 RNA 经过分离提纯之后反转录成 cDNA；3）cDNA 片段化并在一端或者两端加接头修饰；4）加接头序列大量扩增开始并行测序，获得序列信息的初始读长（raw reads）；5）初始读长的质量评估，利用软件如 RobiNA（Lohse *et al.*，2012）去掉低质量、短于设定阈值的读长与读长中的接头序列，获得干净序列（clean reads）进行下游的拼接组装。

4 转录组测序数据组装方法简介

获取尽量完整的转录谱是一切后续分析的基础，因此转录组数据的拼接质量对下游的分析研究至关重要。按照是否需要参考基因组序列，读长的

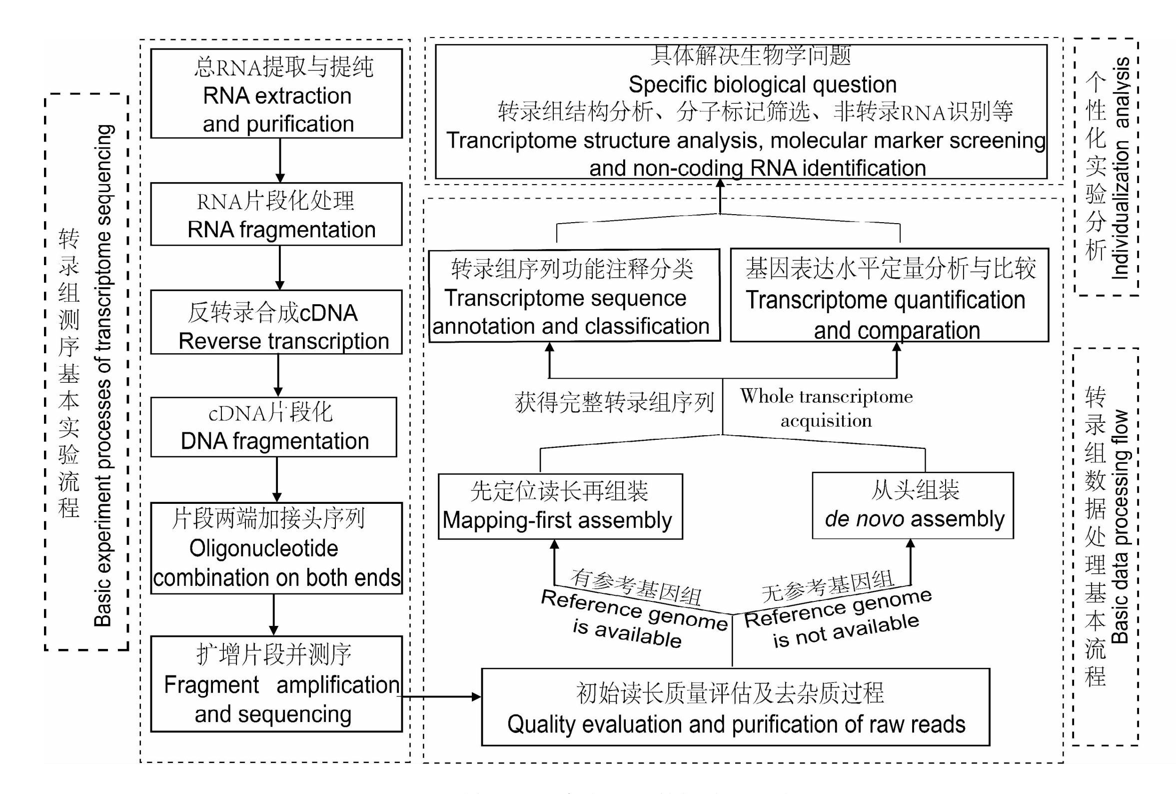


图 1　转录组测序实验和数据分析基本流程

Fig. 1　Basic pipeline of transcriptome sequencing experiments and data analyses

拼接分为两类。第一类为有参考基因组信息的转录组数据拼接,该类方法的主要策略是首先将所有读长比对到参考基因上,再利用读长之间的重叠区(overlap)组装成为完整的 Unigene 序列,并利用一些跨外显子读长的双末端序列信息判断序列中的可变剪接位点。其代表的软件有 Cufflinks(Trapnell *et al.*, 2010),Scripture(Guttman *et al.*, 2010)等。该方法的优势在于识别读长灵敏性强(Grabherr *et al.*, 2011),缺点在于对参考基因组的完整度要求很高。例如 Gan 等(2010)利用 Cufflinks 拼接获得了野生型和 bam 突变型黑腹果蝇 *Drosophila melanogaster* 卵巢和睾丸的转录组,并发现了突变体中大量选择性剪切事件。另外一类为无参考基因的转录组片段组装,又称为“从头(*de novo*)组装”。这种组装方法的原理是在无参考基因组的情况下,首先将读长按照 de Bruijn 图的数据结构分割为多个短语(short words 或 k-mers)(Pevzner *et al.*, 2001),并利用这些短语之间的重叠区直接进行拼接获得 unigene。代表软件为 ABySS(Birol *et al.*, 2009),SOAP *de novo*(Li *et al.*, 2009),Trinity(Grabherr *et al.*, 2011)以及 Velvet(Collins *et al.*, 2008)等,其中 Trinity 在处理短片段的拼接方面有着极佳的表现(Grabherr *et al.*, 2011)。由于摆脱了对基因组序列的依赖,从头拼接广泛应用于非模式物种的研究中,成为了昆虫功能基因组学重要的研究手段之一。例如,Pauchet 等(2009)利用 SOAP *de novo* 拼接法获得了苏云金杆菌感染的白杨叶甲 *Chrysomela tremulae* 中肠转录组;Zhang 等(2013b)利用 Trinity 拼接获得番茄斑萎病毒(TSWV)感染的西花蓟马 *Frankliniella occidentalis* 转录组等。不仅如此,深圳华大基因研究院主持的“千种昆虫转录组进化”项目也是依托其先进的 *de novo* 测序拼接平台对 1 000 种昆虫的转录组进行研究,旨在探索昆虫多样性的原因及其遗传进化机制,了解昆虫与生态系统之间的相互作用,为生物多样性、农业和人类健康等后续研究奠定重要的基础。

5　转录组测序的主要应用

获得完整的转录谱后,转录组数据的下游基本分析一般集中于以下 5 个方面:

1)基因功能注释及获取新转录本。通过将转录本序列在一定的参数下与 nr(non-redundant,非冗余蛋白数据库)、Swiss-Prot、京都基因组百科全书(kyoto encyclopedia of genes and genome, KEGG)和蛋白质直系同源簇(cluster of orthologous groups of

proteins, COG)等数据库比对可获得序列功能信息,同时进行 GO(Gene Ontology,基因的分子功能、生物学过程和细胞组件)(Ashburner *et al.*, 2000)或 COG(Tatusov *et al.*, 2003)等基因功能分类统计,分析基因产物在细胞中的代谢途径(Ogata *et al.*, 1999),可以全面地描述生物体中基因和基因产物的属性,预测 Unigene 功能(Ma *et al.*, 2012),还可以获得低丰度转录本和新转录本(Zhang *et al.*, 2010b)。

2)进行基因转录水平研究(Robinson *et al.*, 2010)。高通量转录组测序可以高灵敏度捕捉不同组织、不同生理状态下的转录组的表达丰度,并用数字基因表达谱(digital gene expression, DGE)监控其动态变化(Audic and Claverie, 1997)。

3)非转录 RNA(non-coding RNA, ncRNA)功能研究,如微 RNA(microRNA)(Dai *et al.*, 2011),小干扰 RNA(small interfering RNA, siRNA)(Lu *et al.*, 2005)以及非编码长 RNA(long non-coding RNA, lncRNA)(Li *et al.*, 2012)等。对这些非编码 RNA 的种类进行鉴定以及其转录水平的测量,有助于从表观遗传学角度了解基因转录与转录后的修饰和调控(Filipowicz *et al.*, 2008)。

4)转录本结构变异研究,即在整个转录组范围内对基因融合(Maher *et al.*, 2009)、RNA 编辑(Graveley *et al.*, 2011)、选择性剪接(Filichkin *et al.*, 2010)等事件进行鉴定识别,全面地了解转录本结构的变异在生物体不同条件下发挥的调控作用(Qi *et al.*, 2011)。

5)分子标记的开发。高通量转录组测序可以为分子标记的开发提供丰富的遗传学资源,主要涉及单核苷酸多态性(single nucleotide polymorphisms, SNPs)和微卫星(simple sequence repeats, SSRs)(Schwarz *et al.*, 2009; Zhang and Yuang, 2013)。

6 转录组测序在鳞翅目昆虫上的应用

Vera 等(2008)发表的利用罗氏 454 测序法获得庆网蛱蝶 *Melitaea cinxia* 转录组的论文标志着将高通量测序法引入鳞翅目物种研究中已经可行。随后,越来越多关于转录组测序在鳞翅目昆虫中应用的研究成果纷纷在各类杂志上发表,所解决的科学问题也慢慢由最初单纯的获得转录组图谱、挖掘重要基因(Zou *et al.*, 2008; Vogel *et al.*, 2011)、而趋向于解决有害物种防治、近缘种遗传关系与进化等更加复杂的实际问题。截止到 2013 年,已有包括直翅目、膜翅目在内的 7 个目的 68 种昆虫进行过转录组测序,发表论文 89 篇(Zhang and Yuang, 2013),极大地促进了昆虫组学的研究。

6.1 转录组测序在鳞翅目昆虫发育研究中的应用

鳞翅目昆虫属于全变态类昆虫,其摄食和移动的能力在发育的不同阶段差异很大。利用高通量测序研究其不同生长阶段的基因表达与调控机制,有助于了解基因转录图谱在整个发育过程中的动态变化,为有害昆虫的防治和益虫保护提供分子依据。例如,Li 等(2013)利用 Illumina 测序平台,获得了亚洲稻类第一害虫稻纵卷叶螟 *Cnaphalocrocis medinalis* 的转录组信息,比较了其 4 个发育时期的转录组,发现数以千计的基因表达量发生了变化,其中卵期与 3 龄幼虫期之间的差异基因数量最多而且多数新陈代谢相关基因在 3 龄幼虫中发生了上调,因此认为该生长阶段具有更强的代谢活力;还发现在不同时期的转录组中,有关蜕皮、DNA 复制以及保幼激素等相关基因多次出现,表明其对发育的关键作用。Gu 等(2013)利用相同的测序手段获得了不同发育阶段斜纹夜蛾 *Spodoptera litura* 的转录组信息,发现近 9 000 个基因的表达量在发育过程中发生变化,其中细胞色素 P450、几丁质酶、丝氨酸蛋白酶和参与性信息素生物合成基因的表达量具有显著改变。该研究还发现与表皮合成息息相关的几丁质酶、b-N-乙酰氨基和几丁质脱乙酰酶在不同发育时期表现出不同的表达模式,前两种物质在预蛹阶段有最高表达,但后者在蛹阶段表达量最高,因此推断几丁质酶和 b-N-乙酰氨基可能负责降解幼虫角质层的退化,而几丁质脱乙酰酶可能有助于降解蛹的角质层。

6.2 转录组测序在鳞翅目昆虫与植物互作研究中的应用

大多数鳞翅目昆虫的幼虫以植物为生,因此一些转录组研究专门针对害虫与寄主之间相互作用的分子机制而展开,旨在了解害虫为害和消化寄主的主要功能基因以及害虫对植物防御毒素的免疫应答机制。斑娥科六星灯蛾 *Zygaena filipendulae* 幼虫可以合成生氰糖苷毒素,主要用于破坏寄主植物百脉根 *Lotus corniculatus* 细胞壁,但是合成该毒素的相关基因和机理却一直不为人知。Zagrobelny 等(2009)利用罗氏 454 测序获得了六星灯蛾的转录组文库并将其与家蚕和黑腹果蝇的基因组相比较,通过序列长度、GC 含量以及重复度等信息确定了其生氰糖苷合成酶的相关基因和代谢通路,研究推断这是双

孔目蛾类一种很原始的特性，来源于与寄主植物的趋同进化。Celorio-Mancera 等（2012）给棉铃虫 *Helicoverpa armigera* 投喂不同的食物（棉花、烟草、菜豆、鹰嘴豆和两种人工食物）以及棉花的不同组织，并检测其中肠基因转录组表达的一系列变化。该研究注释出了大量服用不同食物后显著变化的代谢通路和消化基因。研究发现该虫食用大豆与烟草后中肠转录组反应更相似，而食用棉花花蕾和叶片的中肠基因表达更相似。Fescemyer 等（2013）将产生蛋白酶毒素的转基因玉米喂食秋粘虫 *Spodoptera frugiperda*，并比较了食用有毒玉米和正常玉米的秋粘虫的中肠转录组。研究发现受到毒素伤害的昆虫其编码围食膜组织蛋白的基因出现了明显的上调，可能用于中肠围食膜修复。同时，受毒素伤害的昆虫的消化酶基因家族（肽链内切酶、氨基肽酶、脂肪酶和淀粉酶）也出现了明显的上调，从而维持中肠的正常的能量代谢以补充体能。

6.3　转录组测序在鳞翅目昆虫对杀虫剂抗性方面的应用

鳞翅目幼虫的防治一直是农林业保护的重要课题。化学杀虫剂用量少效果快，但是其大量使用容易使害虫产生不同程度的抗药性。转录组测序可以筛选出大量农药抗性基因，有助于了解害虫抗药性发展和免疫应答机理。例如 He 等（2012）利用 Illumina 测序、Lin 等（2013）用罗氏 454 测序均获得了小菜蛾 *Plutella xyllostella* 的转录组，后者还比较了对氯虫苯甲酰胺抗性能力不同的小菜蛾的转录组表达量。两项研究均鉴定出大量杀虫剂抗性基因，包括细胞色素 P450，谷胱甘肽转移酶（glutathione S-transferase, GST）等基因家族，为小菜蛾抗性研究提供方向。相对于化学杀虫剂而言，微生物源的新型农药对人畜更加安全，对环境也更加友好，因此逐渐成为治理鳞翅目幼虫的主要方法。例如苏云金杆菌 *Bacillus thuringiensis*（Bt）是一种微生物源低毒杀虫剂，该菌的主要活性成分是一种或数种杀虫晶体蛋白（insecticidal crystal proteins, ICPs），又称 δ-内毒素，对鳞翅目幼虫杀伤力巨大，然而其毒理分子机制仍然不为人所知。Ma 等（2012）利用 Illumina 高通量测序获得了二化螟 *Chilo suppressalis* 的中肠细胞转录组，并在前 100 个显著富集的基因中进行筛选做 Western blotting 分析，最终发现氨肽酶的 N 和 EH 域能够与苏云金杆菌毒素 Cry1Ca 结合，揭示了二化螟细胞中识别该毒素的蛋白区域。Sparks 等（2013）用同样的测序平台获得了被苏云金杆菌感染的 舞毒蛾幼虫中肠细胞转录组，并比较了感染前和感染 24 h 后各基因的转录变化。研究表明差异表达的基因主要集中于消化酶类和结合蛋白类，因此推测它们可能是 Bt 毒素攻击害虫最主要的靶位点。防治美洲棉铃虫 *Helicoverpa zea* 方法之一是喷洒棉铃虫单衣壳核多角体病毒。Nguyen 等（2012）利用高通量测序获得了 *H. zea* 的转录组文库并构建成表达微阵列芯片。用该芯片与被棉铃虫单衣壳核多角体病毒感染的细胞转录组进行杂交，比较基因的表达水平在感染后不同时长的变化。研究发现在感染病毒后，其应激反应基因、RNA 代谢基因、酶调节活性基因和高尔基体基因出现了显著的表达变化。研究还发现在感染 18 h 后，四大免疫信号代谢通路（Toll, IMD, Jak-STAT 和 JNK）基因均出现明显的表达下调，表明该时间点为病毒发挥毒理效应的最佳时期。

6.4　转录组测序在鳞翅目分子标记开发上的应用

转录组测序获得的大量基因片段极大地促进了鳞翅目昆虫遗传标记的筛选。豆荚螟 *Maruca vitrata* 是非洲热带地区的主要虫害之一，Margam 等（2011）利用 Illumina 对其肠和唾液腺等组织进行了转录组测序。在获得转录组文库的基础上，该研究筛选出 1 542 个 SNP 位点，其中 71 个位点暗示西非地区的 *M. vitrata* 可能发生了种群遗传结构分化。Pascual 等（2012）用多种病原微生物侵染甜菜夜蛾 *Spodoptera exigua*，并对其中肠、脂肪体等不同组织细胞的转录组进行罗氏 454 和传统测序，获得了潜在的 335 个 SSR 位点 和 17 178 个 SNP 位点。O'Neil等（2010）同样利用罗氏 454 测序法得到弄珠蝶 *Erynnis propertius* 和泽丽凤蝶 *Papilio zelicaon* 两者的转录组，并分别检测到约 3.6 万个和 6.2 万个潜在的 SNP 位点。这些分子标记的开发为生物遗传多样性、基因图谱绘制以及环境保护等研究提供了巨大地帮助。

6.5　转录组测序在鳞翅目昆虫分类与进化方面的应用

单基因数据信息量不足，很难重现复杂的生物进化历史，因此就单纯从数据量上来说，利用转录组测序获得多个基因片段可以消除这种随机误差（Zou and Ge, 2008），更加可靠地推断系统发育关系，为了解鳞翅目的成种机制与进化历程提供帮助。物种间杂交的成种事件在动物界中十分少见，但是凤蝶 *Papilio appalachiensis* 却有可能是 *P. glaucus* 和 *P. canadensis* 两者的后代。为了明确这三者之间的亲缘关系，Zhang 等（2013a）利用 Illumina 测序获得

了以上三者和外群 *P. polytes* 的转录组文库,筛选出来四大类作为适应性进化的基因片段,并且使用系统发育和种群遗传方法,探究三者基因渗入与基因镶嵌的模式。结果表明三者并没有明显的遗传分化,但是 *P. appalachiensis* 在染色体基因上与 *P. glaucus* 更加相似,而在 Z 染色体基因上又与 *P. canadensis* 更加相似,因此倾向于认为该种为杂交种。Ferguson 等(2010)利用罗氏 454 测序技术获得了不同翅型模式的釉蛱蝶的转录组,发现 28 个位于常染色体 HmYb/Sb 区域的等位基因可能与釉蛱蝶翅型模式的决定机制相关,并且发现其中 6 个基因出现过可变剪接事件。研究认为 HmYb/Sb 区域有可能是该蝴蝶翅膀发育和分化的热点区域,因此对研究蝴蝶翅膀的拟态模式、鳞翅目适应辐射和趋同进化有重大意义。

6.6 转录组测序在鳞翅目昆虫非编码 RNA 方面的应用

随着鳞翅目转录组测序研究的深入,研究范围也已经不局限在编码 RNA 上面,而是向着那些更短的,对生命体活动具有重要调节功能的非编码 RNA 上发展。Zhang 等(2012)通过转录组测序,将烟草天蛾 *Manduca sexta* 从胚胎到成虫的 4 个阶段构建了 4 个 miRNA 文库。研究注释出了 163 个保守的 miRNA 以及 13 个稀有 miRNA,还找到 82 个保守的 miRNA 前体出现在多个文库中。研究表明,在该害虫的生长发育过程中,多种 miRNA 的表达量发生了较大的变化,因此推断它们在调控生长发育基因表达中有重要的作用。发展高效 RNAi 介导的病虫害防治技术的第一步是要找到合适的沉默靶基因。为了筛选这一世界性顽固害虫的有效靶基因,Li 等(2013)通过 Illumina 测序获得了甜菜夜蛾 *S. exigua* 转录组,并从中挑选了 9 个候选靶基因设计其特异 siRNA,研究发现包括几丁质酶、血浆谷氨酸羧肽酶等在内的 8 个靶基因在注射 siRNA 后,其 mRNA 丰度有不同程度的减少(20% ~94.3%),还发现虫体对注射不同靶标的 siRNA 出现了发育迟缓、"半蜕皮"现象以及身体黑化等不同的反应。

7 转录组测序所面临的困难与挑战

7.1 测序的"偏好性"

RNA-Seq 以测序为基础,在整个转录组的尺度上可以分辨不同序列单碱基的差别,还能够计算不同样品不同时刻表达量的差别,是了解细胞转录组的一种高通量高精度的方法(Marioni *et al.*, 2008; Bullard *et al.*, 2010)。然而同其他技术一样,转录组测序也有着一些亟待解决的缺点。其中缺点之一就是测序的偏好性问题(Marioni *et al.*, 2008)。测序的偏好性导致了转录组测序获得 reads 的数量不仅取决于基因的表达数量以及该基因在基因组上面重复的次数,还取决于不同基因的 GC 含量、长度,甚至是基因片段所在的位置,这种偏好性使得不同基因之间或者不同细胞之间的转录水平无法直接进行有效的比较。例如 Risso 等(2011)发现过高或者过低 GC 含量的基因片段很容易被忽略,从而导致测序结果结果过低地估计了该类基因的表达量。基因的 GC 含量与基因功能息息相关,因此这种结果会导致对基因组研究的错误认识。Bullard 等(2010)指出越长的基因被测序到的可能性越大,因此会低估很多短基因的表达量。Mortazavi 等(2008)在测序哺乳动物转录组的时候发现,位于转录本 RNA 3′和 5′两端的序列几乎没有被测序。相反,Nagalakshmi 等(2008)在测序酵母 RNA 反转录得到的 cDNA 时,却发现测序强烈地偏好于 cDNA 的片段两端的序列。

7.2 数据分析复杂的标准化过程

测序偏好性普遍出现,打破了当初"比较不同组织不同细胞的转录组不需要复杂的标准化过程"(Mortazavi *et al.*, 2008; Wilhelm *et al.*, 2008)这种过于乐观的估计。不仅如此,不同的测序深度以及转录组本身不同基因重复数量的差异,都会影响到数据分析的抽样过程,因而错误地估计某些片段真正的表达量(Robinson *et al.*, 2010; Robinson and Oshlack, 2010)。这样的情况下就需要一个完善的标准化过程,让两份转录样本能够"公平"地比较转录丰度。例如 Robinson 和 Oshlack(2010)提出了 calcNormFactors 函数,将原始 输入的 RNA 文库大小进行标准化处理,转化成有效 RNA 文库,从而有效地削减 RNA 组成差异所造成的表达量错误估计。Risso 等(2011)开发出的 EDASeq 以及 Hansen 等(2012)开发的 cqn 这两个软件包,致力于降低不同样本中 GC 含量导致的测序偏好性和基因表达量误差,其中 Hansen 等(2012)证明在数据标准化处理后,统计精度提高了 42%。Hansen 等(2012)还利用自然 3 次样条曲线和分位数回归将测序读长数效应引入到了泊松模型中,从而调整基因长度不同所带来的误差。这些复杂的标准化过程的实现需要生物学实验研究人员与生物信息学研究人员的通力协

作，而这无疑增加了转录组测序应用的难度。

7.3　复杂剪接方式增加增加数据组装困难

通过转录组测序获得了高质量的 reads 之后，数据分析的首要步骤就是将这些短片段映射到参考基因组上，或者是利用短片段上的重叠区将其装配成 contig。一般来说，越大的基因组，其复杂度越高。基因组的复杂度表现之一就是基因的反式剪接和可变剪接，这大大增加了信息装配的难度。基因的反式剪接（*trans*-splicing）指的是两条不同的 mRNA 的外显子连接到一起；基因组的可变剪接（alternative splicing）指的是有些基因的一个 mRNA 前体通过不同的剪接方式（选择不同的剪接位点）产生不同的 mRNA 剪接异构体，是调节基因表达和产生蛋白质组多样性的重要机制。这些不同剪接方式的广泛存在使得在序列装配的时候很难识别剪接位点，从而进一步增加了装配的难度。要解决这个问题需要两点：一就是构建剪接位点数据库，将所有已知的剪接方式以及潜在的剪接位点都包含进去（Mortazavi *et al.*, 2008; Wilhelm *et al.*, 2008）；第二就是开发新的计算机算法，准确预测新的拼接方式，能够准确了解两条片段之间可能发生的事件。而无论哪一种方法都需要该物种深厚的遗传学、分子生物学、细胞学的背景知识。

8　小结与展望

尽管存在着上述的困难，转录组测序的广泛应用仍然给昆虫纲第二大目——鳞翅目的研究带来了发展的契机与挑战。无论是辅助开发新基因位点，定量了解代谢模式还是检测免疫应激反应等方面，转录组测序均有很好的发展前景。随着技术的发展，转录组双末端测序、单链特异性测序等方法在测序深度和覆盖度方面都会有更大的突破（Wang *et al.*, 2009）。同时，高速发展的第 3 代单分子实时测序省略扩增步骤，可以在短时间内测得大量低错误率读长片段，从而获得更精确的、信息含量更高的转录组图谱。因此在未来的鳞翅目昆虫研究中，可以扩大所研究的昆虫种类，针对更多的非模式生物、农业重大害虫和经济型昆虫进行转录组测序。随着测序技术成本的不断下降，鳞翅目昆虫转录组数据量必定会大规模增长，而用户友好型软件的开发，会极大地深化人们对整个鳞翅目昆虫转录组结构和内容的认识，而这就方便更加深刻地探索它们的遗传模式、发育方式、调控机制、免疫应答机制、生态间互作方法等问题，从而解决农林害虫防治、经济昆虫保护、鳞翅目昆虫的系统发育甚至及其进化起源等问题。我们希望转录组测序在鳞翅目昆虫研究方面有更加广阔而光明的未来。

参考文献（References）

Ansorge WJ, 2009. Next-generation DNA sequencing techniques. *New Biotechnology*, 25(4): 195－203.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G, 2000. Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1): 25－29.

Audic S, Claverie JM, 1997. The significance of digital gene expression profiles. *Genome Research*, 7(10): 986－995.

Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S, Gerstein M, Snyder M, 2004. Global identification of human transcribed sequences with genome tiling arrays. *Science*, 306(5705): 2242－2246.

Birol I, Jackman SD, Nielsen CB, Qian JQ, Varhol R, Stazyk G, Morin RD, Zhao YJ, Hirst M, Schein JE, Horsman DE, Connors JM, Gascoyne RD, Marra MA, Jones SJ, 2009. *De novo* transcriptome assembly with ABySS. *Bioinformatics*, 25(21): 2872－2877.

Boguski MS, Tolstoshev CM, Bassett DEJr, 1994. Gene discovery in dbEST. *Science*, 265(5181): 1993－1994.

Bullard J, Purdom E, Hansen K, Dudoit S, 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* 11: 94.

Celorio-Mancera MD, Heckel DG, Vogel H, 2012. Transcriptional analysis of physiological pathways in a generalist herbivore: responses to different host plants and plant structures by the cotton bollworm, *Helicoverpa armigera*. *Entomologia Experimentalis et Applicata*, 144(1): 123－133.

Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G, Sementchenko V, Piccolboni A, Bekiranov S, Bailey DK, Ganesh M, Ghosh S, Bell I, Gerhard DS, Gingeras TR, 2005. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science*, 308(5725): 1149－1154.

Cloonan N, Forrest ARR, Kolle G, Gardiner BBA, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G, Robertson AJ, Perkins AC, Bruce SJ, Lee CC, Ranade SS, Peckham HE, Manning JM, McKernan KJ, Grimmond SM, 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods*, 5: 613－619.

Collins LJ, Biggs PJ, Voelckel C, Joly S, 2008. An approach to transcriptome analysis of non-model organisms using short-read sequences. In: Arthur J, Ng SK eds. Genome Informatics 2008, Vol. 21. Imperial College Press, Covent Garden. 3－14.

Dai C, Zhang YM, Zhang Q, Wu ZS, Deng W, Zhang X, Guo KK, Tang QH, Hou B, 2011. A microRNA catalog of swine umbilical vein endothelial cells identified by deep sequencing. *Agricultural*

Sciences in China, 10(9): 1467 – 1474.

David L, Huber W, Granovskaia M, Toedling J, Palm CJ, Bofkin L, Jones T, Davis RW, Steinmetz LM, 2006. A high-resolution map of transcription in the yeast genome. *Proceedings of the National Academy of Sciences of the United States of America*, 103(14): 5320 – 5325.

Emrich SJ, Barbazuk WB, Li L, Schnable PS, 2007. Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Research*, 17: 69 – 73.

Ferguson L, Lee SF, Chamberlain N, Nadeau N, Joron M, Baxter S, Wilkinson P, Papanicolaou A, Kumar S, Kee TJ, Clark R, Davidson C, Glithero R, Beasley H, Vogel H, Ffrench-Constant R, Jiggins C, 2010. Characterization of a hotspot for mimicry: assembly of a butterfly wing transcriptome to genomic sequence at the HmYb/Sb locus. *Molecular Ecology*, 19: 240 – 254.

Fescemyer HW, Sandoya GV, Gill TA, Ozkan S, Marden JH, Luthe DS, 2013. Maize toxin degrades peritrophic matrix proteins and stimulates compensatory transcriptome responses in fall armyworm midgut. *Insect Biochemistry and Molecular Biology*, 43(3): 280 – 291.

Filichkin SA, Priest HD, Givan SA, Shen RK, Bryant DW, Fox SE, Wong WK, Mockler TC, 2010. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Research*, 20(1): 45 – 58.

Filipowicz W, Bhattacharyya SN, Sonenberg N, 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Reviews Genetics*, 9(2): 102 – 114.

Gan QA, Chepelev I, Wei G, Tarayrah L, Cui KR, Zhao KJ, Chen X, 2010. Dynamic regulation of alternative splicing and chromatin structure in *Drosophila* gonads revealed by RNA-seq. *Cell research*, 20(7): 763 – 783.

Glenn TC, 2011. Field guide to next-generation DNA sequencers. *Molecular Ecology Resource*, 11(5): 759 – 769.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 9(7): 644 – 652.

Graveley BR, Brooks AN, Carlson J, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, Brown JB, Cherbas L, Davis CA, Dobin A, Li R, Lin W, Malone JH, Mattiuzzo NR, Miller D, Sturgill D, Tuch BB, Zaleski C, Zhang D, Blanchette M, Dudoit S, Eads B, Green RE, Hammonds A, Jiang L, Kapranov P, Langton L, Perrimon N, Sandler JE, Wan KH, Willingham A, Zhang Y, Zou Y, Andrews J, Bickel PJ, Brenner SE, Brent MR, Cherbas P, Gingeras TR, Hoskins RA, Kaufman TC, Oliver B, Celniker SE, 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature*, 471(7339): 473 – 479.

Gu J, Huang LX, Gong YJ, Zheng SC, Liu L, Huang LH, Feng QL, 2013. *De novo* characterization of transcriptome and gene expression dynamics in epidermis during the larval-pupal metamorphosis of common cutworm. *Insect Biochemistry and Molecular Biology*, 43(9): 794 – 808.

Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, Fan L, Koziol MJ, Gnirke A, Nusbaum C, Rinn JL, Lander ES, Regev A, 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs *Nature Biotechnology*, 28: 503 – 510.

Hansen KD, Irizarry RA, Wu ZJ, 2012. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics*, 13(2): 204 – 216.

Harbers M, Carninci P, 2005. Tag-based approaches for transcriptome research and genome annotation. *Nature Methods*, 2: 495 – 502.

He WY, You MS, Vasseur L, Yang G, Xie M, Cui K, Bai JL, Liu CH, Li XJ, Xu XF, Huang S, 2012. Developmental and insecticide-resistant insights from the *de novo* assembled transcriptome of the diamondback moth, *Plutella xylostella*. *Genomics*, 99(3): 169 – 177.

Huang LL, Cheng TC, Xu PZ, Cheng DJ, Fang T, Xia QY, 2009. A genome-wide survey for host response of silkworm, *Bombyx mori* during pathogen *Bacillus bombyseptieus* infection. *PLoS ONE*, 4(12): e8098.

Li H, Jiang WH, Zhang Z, Xing YR, Li F, 2013. Transcriptome analysis and screening for potential target genes for RNAi-mediated pest control of the beet armyworm, *Spodoptera exigua*. *PLoS ONE*, 8(6): e65931.

Li RQ, Yu C, Li YR, Lam TW, Yiu SM, Kristiansen K, Wang J, 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics*, 25(15): 1966 – 1967.

Li TT, Wang SY, Wu RM, Zhou XY, Zhu DH, Zhang Y, 2012. Identification of long non-protein coding RNAs in chicken skeletal muscle using next generation sequencing. *Genomics*, 99(5): 292 – 298.

Lin QS, Jin FL, Hu ZD, Chen HY, Yin F, Li ZY, Dong XL, Zhang DY, Ren SX, Feng X, 2013. Transcriptome analysis of chlorantraniliprole resistance development in the diamondback moth *Plutella xylostella*. *PLoS ONE*, 8(8): e72314.

Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B, 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, 40(W1): W622 – W627.

Lu C, Tej SS, Luo SJ, Haudenschild CD, Meyers BC, Green PJ, 2005. Elucidation of the small RNA component of the transcriptome. *Science*, 309(5740): 1567 – 1569.

Ma WH, Zhang Z, Peng CH, Wang XP, Li F, Lin YJ, 2012. Exploring the midgut transcriptome and brush border membrane vesicle proteome of the rice stem borer, *Chilo suppressalis* (Walker). *PLoS ONE*, 7(5): e38151.

Maher CA, Palanisamy N, Brenner JC, Cao XH, Kalyana-Sundaram S, Luo SJ, Khrebtukova I, Barrette TR, Grasso C, Yu JD, Lonigro RJ, Schroth G, Kumar-Sinha C, Chinnaiyan AM, 2009. Chimeric transcript discovery by paired-end transcriptome sequencing. *Proceedings of the National Academy of Sciences of the United States*

of America, 106(30): 12353 - 12358.

Margam VM, Coates BS, Bayles DO, Hellmich RL, Agunbiade T, Seufferheld MJ, Sun WL, 2011. Transcriptome sequencing, and rapid development and application of SNP markers for the legume pod borer *Maruca vitrata* (Lepidoptera: Crambidae). *PLoS ONE*, 6(7): e21388.

Marioni J, Mason C, Mane S, Stephens M, Gilad Y, 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research*, 18(9): 1509 - 1517.

Morin RD, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh TJ, McDonald H, Varhol R, 2008. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *BioTechniques*, 45(1): 81 - 94.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B, 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 5: 621 - 628.

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*, 320(5881): 1344 - 1349.

Nguyen Q, Palfreyman RW, Chan LCL, Reid S, Nielsen LK, 2012. Transcriptome sequencing of and microarray development for a *Helicoverpa zea* cell line to investigate *in vitro* insect cell-baculovirus interactions. *PLoS ONE*, 7(5): e36324.

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M, 1999. Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 27(1): 29 - 34.

Okoniewski MJ, Miller CJ, 2006. Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. *BMC Bioinformatics*, 7: 276.

O'Neil ST, Dzurisin JDK, Carmichael RD, Lobo NF, Emrich SJ, Hellmann JJ, 2010. Population-level transcriptome sequencing of nonmodel organisms *Erynnis propertius* and *Papilio zelicaon*. *BMC Genomics*, 11.

Pascual L, Jakubowska AK, Blanca JM, Canizares J, Ferre J, Gloeckner G, Vogel H, Herrero S, 2012. The transcriptome of *Spodoptera exigua* larvae exposed to different types of microbes. *Insect Biochemistry and Molecular Biology*, 42(8): 557 - 570.

Pauchet Y, Wilkinson P, van Munster MV, Augustin S, Pauron D, ffrench-Constant RH, 2009. Pyrosequencing of the midgut transcriptome of the poplar leaf beetle *Chrysomela tremulae* reveals new gene families in Coleoptera. *Insect Biochemistry and Molecular Biology*, 39(5 - 6): 403 - 413.

Pevzner PA, Tang HX, Waterman MS, 2001. An Eulerian path approach to DNA fragment assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 98(17): 9748 - 9753.

Qi YX, Liu YB, Rong WH, 2011. RNA-Seq and its applications: a new technology for transcriptomics. *Hereditas*, 33(11): 1191 - 1202.

Risso D, Schwartz K, Sherlock G, Dudoit S, 2011. GC-content normalization for RNA-Seq data. *BMC Bioinformatics*, 12: 480.

Robinson MD, McCarthy DJ, Smyth GK, 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1): 139 - 140.

Robinson MD, Oshlack A, 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11(3): R25.

Royce TE, Rozowsky JS, Gerstein MB, 2007. Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification. *Nucleic Acids Research*, 35(15): e99.

Schwarz D, Robertson HM, Feder JL, Varala K, Hudson ME, Ragland GJ, Hahn DA, Berlocher SH, 2009. Sympatric ecological speciation meets pyrosequencing: sampling the transcriptome of the apple maggot *Rhagoletis pomonella*. *BMC Genomics*, 10: 633.

Sparks ME, Blackburn MB, Kuhar D, Gundersen-Rindal DE, 2013. Transcriptome of the *Lymantria dispar* (gypsy moth) larval midgut in response to infection by *Bacillus thuringiensis*. *PLoS ONE*, 8(5): e61190.

Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA, 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, 4: 41.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L, 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 28(5): 511 - 515.

Van Belleghem SM, Roelofs D, Van Houdt J, Hendrickx F, 2012. *De novo* transcriptome assembly and SNP discovery in the wing polymorphic salt marsh beetle *Pogonus chalceus* (Coleoptera, Carabidae). *PLoS ONE*, 7(8): e42605.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW, 1995. Serial analysis of gene expression. *Science*, 270(5235): 484 - 487.

Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanksi I, Marden JH, 2008. Rapid transcriptome characterization for a non-model organism using 454 pyrosequencing. *Molecular Ecology*, 17(7): 1636 - 1647.

Vogel H, Altincicek B, Glockner G, Vilcinskas A, 2011. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. *BMC Genomics*, 12: 308.

Wang XW, Luan JB, Li JM, Bao YY, Zhang CX, Liu SS, 2010. De novo characterization of a whitefly transcriptome and analysis of its gene expression during development. *BMC Genomics*, 11: 400.

Wang Z, Gerstein M, Snyder M, 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1): 57 - 63.

Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature*, 453(7199): 1239 - 1243.

Yang F, Shi ZY, Bai SL, Zhang AB, 2013. Comparative studies on species identification of Noctuoidea moths in two nature reserve conservation zones (Beijing, China) using DNA barcodes and thin-film biosensor chips. *Molecular Ecology Resources*, 14(1):

50 – 59.

Yue GD, Gao Q, Luo LH, Wang JY, Xu JH, Yin Y, 2012. The application of high-throughput sequencing technology in plant and animal research. *Scientia Sinica Vitae*, 42(2): 107 – 124. [岳桂东, 高强, 罗龙海, 王军一, 许姣卉, 尹烨, 2012. 高通量测序技术在动植物研究领域中的应用. 中国科学(生命科学版), 42(2): 107 – 124]

Zagrobelny M, Scheibye-Alsing K, Jensen NB, Møller BL, Gorodkin J, Bak S, 2009. 454 pyrosequencing based transcriptome analysis of *Zygaena filipendulae* with focus on genes involved in biosynthesis of cyanogenic glucosides. *BMC Genomics*, 10: 574.

Zhang FJ, Guo HY, Zheng HJ, Zhou T, Zhou YJ, Wang SY, Fang RX, Qian W, Chen XY, 2010a. Massively parallel pyrosequencing-based transcriptome analyses of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV). *BMC Genomics*, 11: 303.

Zhang GJ, Guo GW, Hu XD, Zhang Y, Li QY, Li RQ, Zhuang RH, Lu ZK, He ZQ, Fang XD, Chen L, Tian W, Tao Y, Kristiansen K, Zhang X, Li S, Yang H, Wang J, Wang J, 2010b. Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome Research*, 20(5): 646 – 654.

Zhang QL, Yuang ML, 2013. Progress in insect transcriptomics based on the next-generation sequencing technique. *Acta Entomologica Sinica*, 56(12): 1489 – 1508. [张棋麟, 袁明龙, 2013. 基于新一代测序技术的昆虫转录组学研究进展. 昆虫学报, 56(12): 1489 – 1508]

Zhang W, Kunte K, Kronforst MR, 2013a. Genome-wide characterization of adaptation and speciation in tiger swallowtail butterflies using *de novo* transcriptome assemblies. *Genome Biology and Evolution*, 5(6): 1233 – 1245.

Zhang XF, Zheng Y, Jagadeeswaran G, Ren R, Sunkar R, Jiang HB, 2012. Identification and developmental profiling of conserved and novel microRNAs in *Manduca sexta*. *Insect Biochemistry and Molecular Biology*, 42(6): 381 – 395.

Zhang ZJ, Zhang PJ, Li WD, Zhang JM, Huang F, Yang J, Bei YW, Lu YB, 2013b. *De novo* transcriptome sequencing in *Frankliniella occidentalis* to identify genes involved in plant virus transmission and insecticide resistance. *Genomics*, 101(5): 296 – 305.

Zhu YC, Guo Z, Chen MS, Zhu KY, Liu XF, 2011. Major putative pesticide receptors, detoxification enzymes, and transcriptional profile of the midgut of the tobacco budworm, *Heliothis virescens* (Lepidoptera: Noctuidae). *Journal of Invertebrate Pathology*, 106(2): 296 – 307.

Zou XH, Ge S, 2008. Conflicting gene trees and phylogenomics. *Journal of Systematics and Evolution*, 46(6): 795 – 807. [邹新慧, 葛颂, 2008. 基因树冲突与系统发育基因组学研究. 植物分类学报, 46(6): 795 – 807]

Zou Z, Najar F, Wang Y, Roe B, Jiang HB, 2008. Pyrosequence analysis of expressed sequence tags for *Manduca sexta* hemolymph proteins involved in immune responses. *Insect Biochemistry and Molecular Biology*, 38(6): 677 – 682.

(责任编辑: 袁德成)

更　正

《昆虫学报》2014 年第 57 卷第 6 期第 729 页戚伟平等“节肢动物 ABC 转运蛋白及其介导的杀虫剂抗性”一文中文摘要第 1 句表述有误,按作者意见更正为“腺苷三磷酸结合盒转运蛋白(ATP-binding cassette transporter),简称 ABC 转运蛋白(ABC transporter),是继细胞色素 P450 单加氧酶、羧酸酯酶、谷胱甘肽 S-转移酶之后又一类参与解毒作用的重要蛋白家族,……”;相应地英文摘要第 1 句也需要更正为“ATP-binding cassette transporters (ABC transporters) are a recently documented family of detoxification-involved proteins following cytochrome P450 monooxygenases (P450s), glutathione S-transferases (GSTs) and carboxylesterases (COEs)”。

同期封 2 中文目录中“仿爱夜蛾成虫在莫高窟模拟壁画表面的运动行为及其损害机理”一文作者署名排版错误,应更正为“汪万福,吉爱红,武发思,闫俊峰,赵林毅,张国彬,蔺创业”。由此给作者和读者带来不便,编辑部特此致歉

Erratum

The first sentence of the abstract in the review article “**ATP-binding cassette transporters and their mediated resistance to insecticides in arthropods**” by QI Wei-Ping *et al*. on page 729, *Acta Entomologica Sinica*, 2014, 57(6) was not completely precise. As requested by authors, it should be revised as follows: ATP-binding cassette transporters (ABC transporters) are a recently documented family of detoxification-involved proteins following cytochrome P450 monooxygenases (P450s), glutathione S-transferases (GSTs) and carboxylesterases (COEs). We apologize for any inconvenience caused by this revision.